On a general criterion for nonclassicality

S. Aravinda
Poornaprajna Institute of Scientific Research, Sadashivnagar, Bangalore, India.

R. Srikanth*
*Poornaprajna Institute of Scientific Research, Sadashivnagar, Bangalore, India. and
Raman Research Institute, Sadashivnagar, Bangalore, India.*

The quantum violation of the Leggett-Garg (LG) inequality is less surprising than that of Bell-type inequalities or contextuality inequalities because the latter two involve correlations between commuting observables, and hence do not lead to signaling, whereas the former involves 'signaling in time'/memory. We derive a general Bell-type inequality, encompassing both the spatial and temporal variants, where the classical bound is determined by the maximum violation of the inequality allowed by the signaling in the correlations. This is used to make precise the notion of non-classicality, and in particular, to explain the sense in which some quantum violations of the LG inequality are classical, though sufficiently large quantum violations will be non-classical (thereby demonstrating the quantumness of a qubit). We indicate that this weakened concept of non-classicality also entails features like intrinsic randomness, no-cloning, monogamy of correlations, etc., though to a lesser degree than in non-signaling, nonlocal correlations. We find that quantum invasiveness is, apart from nonlocality and contextuality, a fundamental nonclassical resource in QM, related ultimately to non-commutativity. Quite generally, a theory may be called non-classical if it entails effects that cannot be explained by the signaling available within the theory. Finally, we draw attention to an analogy between non-classicality and the meta-mathematical concept of (Gödel) incompleteness.

arXiv:1211.6407v2 [quant-ph] 4 Dec 2012

Introduction. One of the most remarkable features of quantum mechanics (QM) is nonlocality, characterized by the violation of Bell-type inequalities [1, 2]. Leggett and Garg (LG) [3] proposed the temporal version of a Bell-type inequality that is satisfied by all noninvasive-realist theories (For related variants, cf. Ref. [4]). 'Realism' is the assumption that the given system Q possesses determinate properties prior to measurement [5]. 'Noninvasiveness' is the assumption of measurability of a system without disturbing the subsequent evolution of its possessed value. Thus, measurement only reveals a pre-existing value.

Suppose the quantities $\hat{a}$ or $\hat{a}'$ are measured at time t_A and $\hat{b}$ or $\hat{b}'$ at time $t_B > t_A$, where the observables satisfy $\hat{a}, \hat{a}', \hat{b}, \hat{b}' = \pm 1$. Under the stated assumptions, $\hat{a}(\hat{b}+\hat{b}') + \hat{a}'(\hat{b}-\hat{b}') = \pm 2$, so that

$$|\langle \hat{a}\hat{b}\rangle + \langle \hat{a}\hat{b}'\rangle + \langle \hat{a}'\hat{b}\rangle - \langle \hat{a}'\hat{b}'\rangle| \leq 2, \tag{1}$$

which is the LG inequality. Violation of the LG inequality implies the negation of the assumption of non-invasiveness and/or realism. Note that Eq. (1) is a 2-time (t_A, t_B) variant of the 4-time version that was the original LG inequality, and is essentially the same, except to bring the analogy with the Bell inequality closer [6].

In the microscopic scale, because of non-commutativity of quantum observables, measurement in general 'collapses' the quantum state and is thus invasive. The LG inequality was thus originally proposed for macroscopic systems, which can in principle be measured non-invasively.

The assumptions behind the derivation of the Bell inequality are localism and realism. As a classical theory is necessarily local and realist, a violation of Bell's inequality implies non-classicality. Likewise, as a classical theory is necessarily non-contextual and realist, a violation of a contextuality inequality [7] also implies non-classicality. However, the violation of the LG inequality does not entail non-classicality, since a classical theory is not necessarily non-invasive.

Invasiveness implies a disturbance or signal [8] carried forward in time from one measurement to another on the same particle, such that the probability distribution of a subsequent measurement depends on the choice made earlier [9]. For sufficiently large signal, an invasive-realist classical mechanism can presumably be used to violate the LG inequality. Its violation is thus less surprising than that of Bell-type inequalities, which involve no signaling.

From this perspective, the violation of the LG inequality is not non-classical unless the degree of violation is shown to be larger than can be explained by a classical mechanism that uses the signaling in the correlations. We note that in the context of the LG inequality, 'signaling' means 'signaling in time' or carry-forward of memory. It does not contradict the no-signaling principle that holds between geographically separated particles, and hence is not prohibited by special relativity.

The problem thus arises of clarifying the sense, and scope, in which violation of the LG inequality is nonclassical. More generally, one wishes to extend the notion of non-classicality beyond contexts of commuting observables (as required in a setting appropriate to test a Bell-type or contextuality inequality) to correlations between sequential non-commuting observables, thus determining

the full scope of non-classicality. This is the purpose of this article. Our main result is the derivation of a generalization of the Bell or LG or contextuality inequality which is a more faithful witness of nonclassicality when signaling/memory is present.

Disturbance. A set of correlations $\mathbf{P} \equiv P(x,y|a,b)$ between two parties $\mathcal{A}$ and $\mathcal{B}$ (where x,y and a,b are possible discrete measurement outcomes and settings, respectively) can be described by a joint distribution (JD) if and only if it has a deterministic hidden variable (HV) description $P(x,y|a,b) = \int \rho(\lambda)P(x|a,\lambda)P(y|b,\lambda)d\lambda$ [10]. For the case where $a,b = 0,1$ and $x,y = \pm 1$, this is equivalent to the satisfaction of the inequality $\Lambda(\mathbf{P}) \leq 2$, where $\Lambda(\mathbf{P}) \equiv P(x = y|00) + P(x = y|10) - P(x = y|11) + P(x = y|01) - P(x \neq y|00) - P(x \neq y|10) + P(x \neq y|11) - P(x \neq y|01)$, where $P(x = y|ab) \equiv P(+1,+1|a,b) + P(-1,-1|a,b)$ and $P(x \neq y|ab) \equiv P(+1,-1|a,b) + P(-1,+1|a,b)$. When a and b are sequential measurements on the same particle, then $\Lambda(\mathbf{P})$ is the left-hand side of Eq. (1). When a and b belong to different particles, the above inequality is the usual CHSH inequality [2].

Violation of the CHSH/LG inequality, or equivalently, a lack of JD, constitutes what we call a *disturbance*. Disturbance may or may not be *signaling* (the property that measurement at $\mathcal{A}$ reveals the $\mathcal{B}$ settings or vice versa). In QM, if the random variables denoting measurements, A and B, belong to spatially separated particles (as is the case in a Bell setting), then the disturbance is nonlocal and non-signaling. If a and b pertain to the *same* particle (as in an LG setting), then the disturbance is local and signaling. In local QM, this disturbance arises on account of non-commutativity of observables. For in this case, the product of projectors to these operators $\Pi_{A=\pm1}\Pi_{B=\pm1}$ are not hermitian, and hence do not define a valid probability distribution in terms of their expectation values. Non-signaling disturbance also has its ultimate roots in non-commutativity.

Disturbance implies that the communication cost for simulating $\mathbf{P}$ is greater than 0, i.e., it is impossible for two (or more) players, given pre-shared randomness, to recreate the given correlations using local strategies and without any communication [11]. A lower bound on the average communication cost to simulate $\mathbf{P}$ may be obtained as follows [12]. We may express $\mathbf{P}$ as a convex combination of *deterministic* strategies $\mathbf{P} = \sum_j \sum_{\lambda_j} q_{\lambda_j} \mathbf{d}^{\lambda_j}$, where $\mathbf{d}^{\lambda_j}$ is the λ_j-th deterministic protocol requiring a communication cost of $c_j = 0,1,2,\cdots$ bits to implement. A deterministic protocol is of the form $P(x,y|a,b) \equiv \delta^x_{f(a,b)}\delta^y_{g(a,b)}$, where $f(a,b)$ ($g(a,b)$, resp.) specifies Alice's (Bob's, resp.) outcome measurements a and b (with the convention that in the binary case 0 or 1 outcomes represent ∓ 1). If $f = f(a)$ and $g = g(b)$, then no communication is needed, and this strategy defines local correlations. A convex combination of local strategies clearly defines a classical protocol. If $f = f(a)$ and $g = g(a,b)$ then Alice needs to communicate 1 bit to Bob indicating her choice of a.

For example, the deterministic strategy $\mathbf{S}_0^{a\to b}$ defined by $P(x,y|a,b) = \delta^x_0\delta^y_{a\cdot b}$ has signaling from Alice to Bob, and maximally violates the CHSH inequality. Similarly: $\mathbf{S}_1^{a\to b}$ defined by $P(x,y|a,b) = \delta^x_1\delta^y_{\overline{a\cdot b}}$. The average communication cost of $\mathbf{P}$ is $C(\mathbf{P}) = \min_{\{q_j\}} \sum_j q_j c_j$, where $q_j = \sum_{\lambda_j} q_{\lambda_j}$; that is, the average communication cost, minimized over all possible decompositions of $\mathbf{P}$ into deterministic strategies.

Let $\Lambda_j \equiv \max_{\lambda_j} \Lambda(\mathbf{d}^{\lambda_j})$. One can show that $C(\mathbf{P}) \geq \frac{\Lambda(\mathbf{P})-\Lambda_0}{\Lambda_J-\Lambda_0}c_J$, where J is the integer $j(\neq 0)$ that maximizes $(\Lambda_j - \Lambda_0)/c_j$.

Particularizing to the CHSH or LG inequality, we know that the maximum local bound is given by $\Lambda_0 = 2$. We find that $J = 1$, with $\Lambda_1 = 4$, $c_1 = 1$ bit, so that the above lower bound becomes:

$$C(\mathbf{P}) \geq \frac{1}{2}\Lambda(\mathbf{P}) - 1. \tag{2}$$

For maximal violation of the LG inequality, which is $2\sqrt{2}$, we find the required communication cost $C(\mathbf{P}) \equiv \sqrt{2} - 1 \approx 0.41$. (It may be remarked that Ref. [11] shows that 1 bit of communication suffices to reproduce correlations for an arbitrary pair of projective measurement). In the CHSH case, where $\mathbf{P}$ is non-signaling, the lower bound (2) is also sufficient to reproduce $\mathbf{P}$ [12].

Signaling. In strategy $\mathbf{S}_1^{a\to b}$, the maximum signaling from Alice to Bob (when she chooses $a = 0$ or $a = 1$ with equal probability and Bob always chooses $b = 1$) is also 1 bit. The analogous strategy with signaling from Bob to Alice is $\mathbf{S}_1^{a\leftarrow b}$ given by $\mathbf{P} \equiv \delta^x_{a\cdot b}\delta^y_0$. Suppose the correlations $\mathbf{P}$ contain signaling from Alice to Bob (but not the reverse), then we can interpret the correlations temporally with Bob measuring the state measured by Alice. Provided the correlations are interpreted locally (as it is in the LG situation), signaling represents memory capacity and then there is no contradiction with the no-signaling principle or special relativity.

Let S_j denote the maximum signaling allowed by any j-bit one-way signaling strategy. We may regard Alice's input random variable A, and Bob's output random variable Y as constituting a Markov chain $A \to Y$ [13]. By the data processing inequality, $c_j = H(A) \geq I_{\max}(A:Y) = S_j$, where $H(A)$ is the Shannon entropy and $I_{\max}(A:Y)$ denotes the maximum mutual information $I(A:Y) = H(A) - H(A|Y)$ [13]. By convexity, a probabilistic protocol over deterministic strategies with at most c_j bits of communication cost will entail signaling $S(\mathbf{P})$ of a most c_j bits, and so

$$S(\mathbf{P}) \leq C(\mathbf{P}). \tag{3}$$

In particular, a uniform mixture of $\mathbf{S}_0^{a\to b}$ and $\mathbf{S}_1^{a\to b}$ yields the PR box [14], which is non-signaling but maximally violates the CHSH inequality.

Specializing to the case of two measurements and two outcomes, given correlations $\mathbf{P}$, let P_0^b (P_1^b, resp.) represent the probability that Bob, measuring observable b, finds $y=+1$ when Alice measures observable $a=0$ ($a=1$, resp.). The necessary condition for signaling is $s \equiv \max_b |P_0^b - P_1^b| > 0$, where b. The signalled information $S(\mathbf{P})$ received by Bob is quantified by the mutual information $I(A:Y)$ maximized over Alice's and Bob's choices. Letting α and $\beta \equiv 1-\alpha$ denote the probabilities with which Alice chooses $\hat{a}$ or $\hat{a}'$, $Q_j^b \equiv 1 - P_j^b$, $\overline{P} \equiv \alpha P_0^b + \beta P_1^b$ and $\overline{Q} \equiv \alpha Q_0^b + \beta Q_1^b$, explicitly

$$S(\mathbf{P}) = \max_{b,\alpha} \left[H(\alpha) + \alpha P_0^b \log\left(\alpha P_0^b/\overline{P}\right) + \beta P_1^b \log\left(\beta P_1^b/\overline{P}\right) + \alpha Q_0^b \log\left(\alpha Q_0^b/\overline{Q}\right) + \beta Q_1^b \log\left(\beta Q_1^b/\overline{Q}\right)\right]. \tag{4}$$

The above may straightforwardly be generalized when Alice and Bob have more than 2 settings. It may be verified that as expected, $S(\mathbf{P}) = 0$ when $P_0^{\underline{b}} = P_1^{\underline{b}}$, and is, maximally, $S(\mathbf{P}) = 1$ when $P_0^{\underline{b}} = 0$ and $P_1^{\underline{b}} = 1$ or vice versa, where $\underline{b}$ is the value of b that maximizes $S(\mathbf{P})$.

For a qubit in QM, the correlations for sequential measurements are given by

$$P(x,y|\hat{a},\hat{b}) \ = \ \mathrm{Tr}\left(\frac{1+y\hat{b}}{2}\frac{1+x\hat{a}}{2}\rho\frac{1+x\hat{a}}{2}\right) = \frac{1}{4} + \frac{x}{4}\mathrm{Tr}(\hat{a}\rho) + \frac{y}{8}\mathrm{Tr}(\hat{b}\rho) + \frac{xy}{8}\mathrm{Tr}(\{\hat{a},\hat{b}\}\rho) + \frac{y}{8}\mathrm{Tr}(\hat{a}\hat{b}\hat{a}\rho). \tag{5}$$

Without loss of generality, let $\underline{b} = \hat{b}$. From Eq. (5), the condition for signaling is

$$s \equiv |P_0^{\hat{b}} - P_1^{\hat{b}}| = \left|\frac{1}{4}\mathrm{Tr}\left[\left(\hat{a}\hat{b}\hat{a} - \hat{a}'\hat{b}\hat{a}'\right)\rho\right]\right| > 0, \tag{6}$$

which is bounded above in QM as: $|s| \le \frac{1}{4}\left|\mathrm{Tr}\left(\hat{a}\hat{b}\hat{a}\rho\right)\right| + \frac{1}{4}\left|\mathrm{Tr}\left(\hat{a}'\hat{b}\hat{a}'\rho\right)\right| \le \frac{1}{2}$. Let the density operator obtained by measuring $\hat{a}$ (resp., $\hat{a}'$) on $|\psi\rangle$ be ρ_0 (resp., ρ_1). The above bound on signaling strength s can also be stated in terms of trace distance: $\tau \equiv \frac{1}{2}||\rho_0 - \rho_1|| > 0$. Intuitively, we expect to maximize s in Eq. (6) or τ when $|\psi\rangle$ is an eigenstate of $\hat{a}$, and $\hat{a}'$ maximally fails to commute with $\hat{a}$ (i.e., the two observables form a pair of *mutually unbiased bases*). By a similar argument as used to bound s, trace distance satisfies $\frac{1}{2}||\rho_0 - \rho_1|| \le \frac{1}{2}$. This inequality, as well as (6), are saturated for the settings:

$$|\psi\rangle = |0\rangle; \hat{a} = \sigma_z, \hat{a}' = \sigma_x; \hat{b} = \sigma_z. \tag{7}$$

It follows from Eq. (6), $[\hat{a},\hat{b}] \ne 0$ and/or $[\hat{a}',\hat{b}] \ne 0$ is necessary for signaling, or equivalently, an interference between the eigenstates of the operators $\hat{a}$ and $\hat{a}'$ [9]. If Alice and Bob are geographically separated, and ρ is a two-qubit state, Eq. (5) still holds, with the representations $\hat{a} \leftarrow \hat{a} \otimes \mathbb{I}$, $\hat{a}' \leftarrow \hat{a}' \otimes \mathbb{I}$, $\hat{b} \leftarrow \mathbb{I} \otimes \hat{b}$, $\hat{b}' \leftarrow \mathbb{I} \otimes \hat{b}'$. We have $[\hat{a},\hat{b}] = [\hat{a}',\hat{b}] = 0$, and we obtain the usual no-signaling. It may be noted that, even in the local case, there is no backward signaling from Bob to Alice, i.e., Alice's outcome probabilties as derived from Eq. (5) are independent of Bob's settings, as expected.

From Eq. (5), we obtain the correlator $\langle\hat{a}\hat{b}\rangle = \sum_{x,y} xyP(x,y) = \frac{1}{2}\langle\{\hat{a}\hat{b}\}\rangle = \vec{a}\cdot\vec{b}$, where $\hat{a} = \vec{a}\cdot\vec{\sigma}$ and $\hat{b} = \vec{b}\cdot\vec{\sigma}$. We note that the correlator is independent of temporal order, at least for qubits, even though there is a signaling from Alice to Bob.

Thus, the correlations are the same as that obtained by von Neumann measurements on singlets, and it follows that the LG inequality is microscopically violated, and that the Cirelson bound [15] for Bell-type inequalities also holds for the LG inequality (1) [8].

If the state is the maximally mixed $I/2$, then $|P_0^b - P_1^b| \propto \mathrm{Tr}\left(\hat{a}\hat{b}\hat{a} - \hat{a}'\hat{b}\hat{a}'\right) = 0$, since $\hat{a}^2 = (\hat{a}')^2 = I$ for any qubit observable with spectrum ± 1. Similarly $\tau = 0$. Interestingly, because the correlators $\langle\hat{a}\hat{b}\rangle$ are state-independent, this state will nevertheless violate the LG inequality maximally.

Main result: Non-classicality. In the classical world, clearly Alice and Bob can violate Bell-type inequalities provided she can signal to him. Intuitively, $\mathbf{P}$ is classical if it allows signaling strong enough to accommodate the inequality violation it entails. A deterministic strategy like $\mathbf{S}_1^{a\to b}$ is classical in the sense that the maximum signaled information allowed in it, which is 1 bit, can be used to augment a local strategy to re-produce the disturbance property of $\mathbf{S}_1^{a\to b}$, such as maximally violating

the CHSH inequality.

(1) General criterion: If the maximum signalable information in $\mathbf{P}$ exceeds the minimum communication cost for reproducing the disturbance, then clearly the inequality violation can be simulated classically given the same amount of signaling. (We don't require that Nature should use the same strategies.) Let $C_\Lambda(\mathbf{P}) \equiv \frac{1}{2}\Lambda(\mathbf{P})-1$. Note that $C(\mathbf{P}) = C_\Lambda(\mathbf{P})$ for non-signaling $\mathbf{P}$, and in general $C(\mathbf{P}) \geq C_\Lambda(\mathbf{P})$, as follows from Eq. (2). From this perspective, the necessary and sufficient condition for the classicality of $\mathbf{P}$, even in the presence of signaling, is:

$$S(\mathbf{P}) \geq C_\Lambda(\mathbf{P}). \tag{8}$$

Eq. (8) generalizes both the Bell and LG inequalities to give a general criterion for classicality.

(2) Bell, LG and contextuality inequalities in QM: In the spatial (i.e., usual) Bell inequality scenario, by no-signaling $S(\mathbf{P}) = 0$, so that Eq. (8) yields $C_\Lambda(\mathbf{P}) \leq 0$, which is just the usual CHSH inequality. By the criterion (8), any violation of the CHSH inequality is thus necessarily non-classical, as expected.

For sequential measurements of a qubit, $S(\mathbf{P})$ in Eq. (4) is bounded above by the Holevo quantity [13] $\chi = S(\alpha\rho_0 + \beta\rho_1) - \alpha S(\rho_0) - \beta S(\rho_1)$. Direct substitution using settings (7) in (4) yields $H(\alpha) + \alpha \log\left(\frac{2\alpha}{1+\alpha}\right) + \frac{1-\alpha}{2}\log\left(\frac{1-\alpha}{1+\alpha}\right)$, which, when maximized, yields $S(\mathbf{P}) = \log(5) - 2 \approx 0.32$ bits at $\alpha = 3/5$. This is the maximum signaling quantumly possible from Alice to Bob [8], and saturates the Holevo bound. Since the minimum communication cost for a maximal quantum violation (of $2\sqrt{2}$) of the LG inequality is about 0.41 bits, violation of the inequality (8) follows. Thus maximal *quantum* violation of the LG inequality is indeed non-classical: a communication cost of at least $0.41 - 0.32 = 0.09$ bits is required on average over the signaling provided by the invasiveness, in order to violate the LG inequality in a classical simulation.

On the other hand, non-maximal violations of the LG inequality are not necessarily non-classical, as seen from Figure 1, where $\hat{a}, \hat{b}, \hat{a}'$ and $\hat{b}'$ are oriented in the xz plane, and separated by angular intervals θ; the initial state is the $+1$ eigenstate of $\hat{a}$.

Setting $C_\Lambda(\mathbf{P}) \leq \log(5) - 2$, we find that $\Lambda(\mathbf{P}) \leq 2(\log(5) - 1) \approx 2.64$, a tighter version of the LG inequality, that gives a sufficient condition for nonclassicality. That it is not necessary is clear from Figure 1, in the range approximately given by $\theta \in [0.99, 1.03]$.

Eq. (8) suggests that one can define a measure of non-classicality, the normalized communication cost : $\eta \equiv \min\{0, (C_\Lambda(\mathbf{P}) - S(\mathbf{P}))/C_\Lambda(\mathbf{P})\}$, with the convention $\eta = 0$ if $C_\Lambda(\mathbf{P}) = 0$. For correlations that are classical, $\eta = 0$. Note that the maximally nonlocal and signaling correlations $\mathbf{S}_1^{a\to b}$ and $\mathbf{S}_1^{a\leftarrow b}$ are classical because signaling can precisely account for the nonlocality.

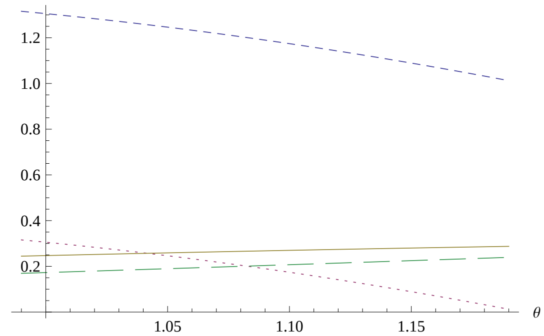

FIG. 1: For the range of angles θ given, the LG inequality is violated, with normalized LG inequality violation ($\Lambda(\mathbf{P})/\Lambda_0$, small-dashed line) exceeding 1. In the region where signaling ($S(\mathbf{P})$ with maximization over b restricted to $\{\hat{b}, \hat{b}'\}$, large-dashed line) exceeds the communication cost (C_Λ, dotted line), above $\theta \gtrsim 1.08$, the LG inequality violations in QM are *not* non-classical, according to Eq. (8). The plain line marks the Holevo bound (asymptotically accessible signal).

For any non-signaling violation of the CHSH inequality, $\eta = 1$, indicating maximal nonclassicality; while for the maximal quantum violation of the LG inequality, we find $\eta \approx 0.09/0.41 \approx 0.22$. In this sense, the maximal quantum violation of the CHSH inequality is *more* nonclassical than the maximal quantum violation of the LG inequality. Interestingly, as we saw earlier, for the maximally mixed state $S(\mathbf{P}) = 0$, implying $\eta = 1$. Thus the maximal violation of the LG inequality by the maximally mixed state is more non-classical than that by a pure state. In contextuality inequalities [7], $C_\Lambda(\mathbf{P}) > 0$ in order to supply context information, but $S(\mathbf{P}) = 0$ because observables in a correlator are pairwise commutative, so that correlations are maximally non-classical.

Intuitively, a single qubit is a non-classical object. However, Bell has shown that the outcome of a single projective measurement on any qubit state can be classically simulated with a hidden variable (HV) model [16]. Further, a proof of non-classicality via contextuality requires dimension greater than 2, which thus does not apply to a qubit subjected to projective measurements. Our result above shows that correlations based on *sequential* measurements on a qubit are nonclassical, thereby making a qubit nonclassical.

Instrinsic randomness, monogamy, etc. The above results apply not just to QM but to any theory in which correlations $\mathbf{P}$ can be defined. A theory is classical if $\forall_{\mathbf{P}}(\eta) = 0$, while it is nonclassical if $\exists_{\mathbf{P}}(\eta) > 0$. It is known that fundamental randomness, Heisenberg uncertainty, privacy and monogamy of correlations, and the impossibility of perfect cloning are necessary features in non-signaling nonlocal theories [17]. In theories of intermediate nonclassicality (i.e., $0 < \eta < 1$), we expect these features to diminish in strength but not vanish. We briefly illustrate this point here. For example, vio-

lation of the LG inequality by a qubit can be used to perform a computation more efficiently than a classical bit [6]. Suppose Alice and Bob share a PR box defined by $a \cdot b = x \oplus y$. With perfect cloning by Bob, one has $a \cdot b' = x \oplus y'$, so that: $a \cdot (b \oplus b') = y \oplus y'$, from which Bob determininstically obtains Alice's input. More generally, if he can clone it only with probability p (i.e., his clone is a mixture of a PR box with probability p and a random outcome state), the signal received by Bob is seen to $I(A:Y) = 1 - H((1-p)/2.)$, a monotonic function in the range $p \in [0,1]$, indicating that with greater bits signaled, there is lesser no-cloning.

Consider the correlation $\mathbf{Q}(p) = p\mathbf{S}_0^{a \rightarrow b} + (1-p)\mathbf{S}_1^{a \rightarrow b}$. Defining intrinsic indeterminancy $I \equiv \min\{p, 1-p\}$ and signal s by Eq. (6), we find $s + 2I = 1$, implying a trade-off between signaling and randomness (More generally, $s + 2I \geq 1$ [18]). Since $C_\Lambda(\mathbf{Q}) = 1$, and $s = 1 \Leftrightarrow S(\mathbf{Q}) = 1$, by our definition, $s < 1$ indicates non-classicality, which is equivalent to $I > 0$. Thus, our signaling-based definition of non-classicality is consistent with the traditional notion associating non-classicality with intrinsic randomness, etc.

As noted earlier, the disturbance in the local sector is an obvious consequence of non-commutativity. The signaling also is a consequence of non-commutativity, as follows from Eq. (6). Thus the nonclassicality of quantum invasive measurements, like that of nonlocality and contextuality, also follows ultimately from quantum non-commutativity. It is different from them in that it involves correlators between non-commuting observables. This shows that quantum invasiveness should be thought of as another basic nonclassical feature, that is at par with contextualtiy and nonlocality.

Discussions and Conclusions. We now mention briefly the similarity of our main result to a theorem in meta-mathematics, the study of mathematics using mathematical tools. (More generally, a *meta-theory* is a theory about a theory). At the moment, the analogy is admittedly very sketchy, but we feel that it offers potential clarity in the way we think about physical theories by permitting an 'outside view'. A well-known metamathematical result is that of incompleteness due to Gödel [19], according to which any axiomatization $\mathbf{A}$ of arithmetic, if it is consistent, then it is *incomplete*, in the sense of there being truths (theorems) expressible in $\mathbf{A}$ but not provable within $\mathbf{A}$. An existential proof of the result is that the set of theorems in $\mathbf{A}$ has the cardinality of the continuum, i.e., it is uncountably large, whilst the number of proofs is only countably infinite.

To be precise, the concept of communication cost is meta-theoretic, as therefore are Eqs. (3) and (8), which are statements *about* QM. Given theory $\mathbb{T}$, predictions in it are, technically, theorems, while signaling indicates a sequence of causes and effects, which is like a train of logical inferences, and thus constitute a proof. Completeness entails that all predicted effects in the theory are explainable via manifest signaling. Consistency entails that only predicted effects are explainable via signaling. Thus non-classicality corresponds to incompleteness in that there exist effects not attributable to signals. A little thought shows that it is related to EPR incompleteness [5].

The connection to Gödel's result is more apparent in its information theoretic version, due to Chaitin [20], according to which, a theorem cannot be derived from axioms that contain less information than the theorem. In the present context, nonclassicality is analogously the case when a theory (a set of theorems) cannot be explained from initial conditions (axioms) and given rules of inference (the available signaling). This suggests that nonclassicality is not pathological but a natural state of affairs in 'theoryspace'.

Acknowledgments. SA acknowledges support through the INSPIRE fellowship [IF120025] by the Department of Science and Technology, Govt. of India.

REFERENCES

* Electronic address: srik@poornaprajna.org

[1] J. Bell, Physics **1**, 195 (1964).

[2] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[3] A. J. Leggett and A. Garg, Phys. Rev. Lett. **54**, 857 (1985).

[4] R. Lapiedra, Europhys. Lett. **75**, 202 (2006).

[5] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[6] C. Brukner, S. Taylor, S. Cheung, and V. Vedral, arXiv:quant-ph/0402127v1.

[7] P. Badziag, I. Bengtsson, A. Cabello, and I. Pitowsky, Phys. Rev. Lett. **103**, 050401 (2009).

[8] T. Fritz, New Journal of Physics **12**, 083055 (2010).

[9] J. Kofler and C. Brukner, arXiv:quant-ph/1207.3666.

[10] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[11] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).

[12] S. Pironio, Phys. Rev. A **68**, 062102 (2003).

[13] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, 2004), 1st ed., ISBN 0521635039.

[14] S. Popescu and D. Rohrlich, Foundations of Physics **24**, 379 (1994).

[15] B. S. Cirelson, Lett. Math. Phys. **4**, 93 (1980).

[16] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[17] L. Masanes, A. Acin, and N. Gisin, Phys. Rev. A **73**, 012112 (2006).

[18] M. J. W. Hall, Phys. Rev. A **82**, 062117 (2010).

[19] K. Gödel, On Formally Undecidable Propositions of Principia Mathematica and Related Systems (Dover, 1962).

[20] G. Chaitin, Int. J. of Theor. Phys. **21**, 941 (1982).